AFRL-OSR-VA-TR-2013-0620

ALGORITHMS FOR LEARNING AND DECISION MAKING

DIMITRI BERTSEKAS

MASSACHUSETTS INSTITUTE OF

12/09/2013
Final Report

DISTRIBUTION A: Distribution approved for public release.

AIR FORCE RESEARCH LABORATORY
AF OFFICE OF SCIENTIFIC RESEARCH (AFOSR)/RSL
ARLINGTON, VIRGINIA 22203
AIR FORCE MATERIEL COMMAND

REPORT DOCUMENTATION PAGE

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*

4. TITLE AND SUBTITLE

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(include area code)*

INSTRUCTIONS FOR COMPLETING SF 298

1. REPORT DATE. Full publication date, including day, month, if available. Must cite at least the year and be Year 2000 compliant, e.g. 30-06-1998; xx-06-1998; xx-xx-1998.

2. REPORT TYPE. State the type of report, such as final, technical, interim, memorandum, master's thesis, progress, quarterly, research, special, group study, etc.

3. DATES COVERED. Indicate the time during which the work was performed and the report was written, e.g., Jun 1997 - Jun 1998; 1-10 Jun 1996; May - Nov 1998; Nov 1998.

4. TITLE. Enter title and subtitle with volume number and part number, if applicable. On classified documents, enter the title classification in parentheses.

5a. CONTRACT NUMBER. Enter all contract numbers as they appear in the report, e.g. F33615-86-C-5169.

5b. GRANT NUMBER. Enter all grant numbers as they appear in the report, e.g. AFOSR-82-1234.

5c. PROGRAM ELEMENT NUMBER. Enter all program element numbers as they appear in the report, e.g. 61101A.

5d. PROJECT NUMBER. Enter all project numbers as they appear in the report, e.g. 1F665702D1257; ILIR.

5e. TASK NUMBER. Enter all task numbers as they appear in the report, e.g. 05; RF0330201; T4112.

5f. WORK UNIT NUMBER. Enter all work unit numbers as they appear in the report, e.g. 001; AFAPL30480105.

6. AUTHOR(S). Enter name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. The form of entry is the last name, first name, middle initial, and additional qualifiers separated by commas, e.g. Smith, Richard, J, Jr.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES). Self-explanatory.

8. PERFORMING ORGANIZATION REPORT NUMBER. Enter all unique alphanumeric report numbers assigned by the performing organization, e.g. BRL-1234; AFWL-TR-85-4017-Vol-21-PT-2.

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES). Enter the name and address of the organization(s) financially responsible for and monitoring the work.

10. SPONSOR/MONITOR'S ACRONYM(S). Enter, if available, e.g. BRL, ARDEC, NADC.

11. SPONSOR/MONITOR'S REPORT NUMBER(S). Enter report number as assigned by the sponsoring/ monitoring agency, if available, e.g. BRL-TR-829; -215.

12. DISTRIBUTION/AVAILABILITY STATEMENT. Use agency-mandated availability statements to indicate the public availability or distribution limitations of the report. If additional limitations/ restrictions or special markings are indicated, follow agency authorization procedures, e.g. RD/FRD, PROPIN, ITAR, etc. Include copyright information.

13. SUPPLEMENTARY NOTES. Enter information not included elsewhere such as: prepared in cooperation with; translation of; report supersedes; old edition number, etc.

14. ABSTRACT. A brief (approximately 200 words) factual summary of the most significant information.

15. SUBJECT TERMS. Key words or phrases identifying major concepts in the report.

16. SECURITY CLASSIFICATION. Enter security classification in accordance with security classification regulations, e.g. U, C, S, etc. If this form contains classified information, stamp classification level on the top and bottom of this page.

17. LIMITATION OF ABSTRACT. This block must be completed to assign a distribution limitation to the abstract. Enter UU (Unclassified Unlimited) or SAR (Same as Report). An entry in this block is necessary if the abstract is to be limited.

Final Report, GRANT NO FA9550-10-1-0412

ALGORITHMS FOR LEARNING AND DECISION MAKING

PI: Prof. Dimitri P. Bertsekas, Massachusetts Institute of Technology

We have investigated learning algorithms for inference and decision making, by using exact and approximate optimization methods. Most of our research has been in approximate dynamic programming/reinforcement learning methods, with a focus on Markovian Decision Problems with a very large number of states. Much of our work is related to a fundamental algorithm, Q-learning, and related new methods that relate to exact and approximate policy iteration. In particular, we have investigated, convergence issues, error bounds, policy oscillation, exploration-enhanced methods, and issues of decision making in a multi-agent environment.

Another research area is large-scale convex optimization methods, with a focus on problems whose cost function involves a sum of a large number of component functions. This includes a unifying framework for polyhedral approximation recently proposed by the principal investigator, incremental gradient and subgradient methods, which are currently at the forefront of algorithmic machine learning research, as well as a new incremental version of the proximal minimization algorithm. We have developed new polyhedral approximation algorithms, including a simplicial decomposition method that applies to large-scale conic programming problems.

Complete List of Publications Supported in Part by this Grant:

D. P. Bertsekas, "Approximate Policy Iteration: A Survey and Some New Methods," Lab. for Information and Decision Systems Report LIDS-P-2833, MIT, April 2010; Journal of Control Theory and Applications, Vol. 9, pp. 310-335.

D. P. Bertsekas, "Centralized and Distributed Newton Methods for Network Optimization and Extensions," Lab. for Information and Decision Systems Report LIDS-P-2866, MIT, April 2011.

D. P. Bertsekas and H. Yu, "Distributed Asynchronous Policy Iteration in Dynamic Programming," Proc. of 2010 Allerton Conference on Communication, Control, and Computing, Allerton Park, ILL, Sept. 2010.

D. P. Bertsekas and H. Yu, "Q-Learning and Enhanced Policy Iteration in Discounted Dynamic Programming," Lab. for Information and Decision Systems Report LIDS-P-2831, MIT, April, 2010; Math. of Operations Research, Vol. 37, 2012, pp. 66-94.

D. P. Bertsekas and H. Yu, "A Unifying Polyhedral Approximation Framework for Convex Optimization," SIAM J. on Optimization, Vol. 21, 2011, pp. 333-360.

D. P. Bertsekas, "Incremental Proximal Methods for Large Scale Convex Optimization," Lab. for Information and Decision Systems Report LIDS-P-2847, MIT, August 2010; Mathematical Programming, Vol. 129, 2011, pp.163-195.

H. Yu and D. P. Bertsekas, "On Boundedness of Q-Learning Iterates for Stochastic Shortest Path

Problems," Lab. for Information and Decision Systems Report LIDS-P-2859, MIT, March 2011; Mathematics of Operations Research 38(2), pp. 209-227, 2013.

 H. Yu and D. P. Bertsekas, "A Mixed Value and Policy Iteration Method for Stochastic Control with Universally Measurable Policies," Lab. for Information and Decision Systems Report LIDS-P-2905, MIT, July 2013.

H. Yu and D. P. Bertsekas, "Weighted Bellman Equations and their Applications in Dynamic Programming," Lab. for Information and Decision Systems Report LIDS-P-2876, MIT, October 2012.

D. P. Bertsekas, "Weighted Sup-Norm Contractions in Dynamic Programming: A Review and Some New Applications", Lab. for Information and Decision Systems Report LIDS-P-2884, MIT, May 2012.

M. Wang and D. P. Bertsekas, "Stabilization of Stochastic Iterative Methods for Singular and Nearly Singular Linear Systems", Lab. for Information and Decision Systems Report LIDS-P-2878, MIT, December 2011 (revised March 2012); to appear in Mathematics of Operations Research.

M. Wang and D. P. Bertsekas, "Convergence of Iterative Simulation-Based Methods for Singular Linear Systems", Lab. for Information and Decision Systems Report LIDS-P-2879, MIT, December 2011 (revised April 2012); to appear in Stochastic Systems.

M. Wang and D. P. Bertsekas, "Incremental Constraint Projection-Proximal Methods for Nonsmooth Convex Optimization ", Lab. for Information and Decision Systems Report LIDS-P-2907, MIT, July 2013.

M. Wang and D. P. Bertsekas, "Incremental Constraint Projection Methods for Variational Inequalities", Lab. for Information and Decision Systems Report LIDS-P-2898, MIT, December 2012.

D. P. Bertsekas, "Lambda-Policy Iteration: A Review and a New Implementation", Lab. for Information and Decision Systems Report LIDS-P-2874, MIT, October 2011. To appear in "Reinforcement Learning and Approximate Dynamic Programming for Feedback Control," by F. Lewis and D. Liu (eds.), IEEE Press Computational Intelligence Series.

 D. P. Bertsekas and H. Yu, "Stochastic Shortest Path Problems Under Weak Conditions," Lab. for Information and Decision Systems Report LIDS-P-2909, MIT, August 2013.

H. Yu and D. P. Bertsekas, "A Mixed Value and Policy Iteration Method for Stochastic Control with Universally Measurable Policies," Lab. for Information and Decision Systems Report LIDS-P-2905, MIT, July 2013.